



S/N 08/720,373

130-126

2308

#6

PATENT
RECEIVED

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APR 17 1997

Applicant: John C. Mitchell et al.

Examiner: Unknown

GROUP 2300

Serial No.: 08/720,373

Group Art Unit: Unknown

Filed: September 27, 1996

Docket: 491.003US1

RECEIVED

Title: DATA PROCESSING METHOD AND APPARATUS

JUN 21 1997

DIRECTOR'S OFFICE
GROUP 2300

PETITION TO MAKE SPECIAL UNDER 37 C.F.R. § 1.102(d)

Assistant Commissioner for Patents
Washington, D.C. 20231

Dear Sir:

This is a Petition to Make Special for the above-identified patent application for advancement of its examination under 37 C.F.R. § 1.102(d). The petition fee of \$130.00 as set forth in § 1.17(i)(2) which is required pursuant to 37 C.F.R. § 1.102(d) is enclosed.

The grounds and conditions for granting this application's special status for advance examination are found in M.P.E.P § 708.02 VIII entitled "Special Examining Procedure for Certain New Applications -- Accelerated Examination." In furtherance of the submittal of this Petition to Make Special accompanied by the fee set forth in 37 C.F.R. § 1.17(i)(2), the Applicants submit the following:

- (a) A preexamination search was conducted by the United Kingdom Patent Office.

The resulting references are listed on the attached United Kingdom Patent Office Search and Examination Reports. The references include:

U.S. Patents

Patent No.

Patentee

Issue Date

5,031,113

Höllerbauer

July 9, 1991

250 TL 04/11/97 08720373
1 126 130.00 CK

Adjustment Date: 07/24/1997 LCHALMER Nm/Nb: 08720373
04/11/1997 250TL 0000011901 FC:126 -130.00 OP

07/24/1997 LCHALMER 00000035 08720373
130.00 OP
01 FC:122

Foreign Patents

<u>Patent No.</u>	<u>Patentee</u>	<u>Publication Date</u>
EP 0649144 A1	IBM	April 19, 1995
EP 0077194 A1	Sharp	April 20, 1983
WO 93/07562 A1	Riverrun	April 15, 1993
GB 2088106 A	Marconi	June 3, 1982
GB 2230370 A	Smiths	October 17, 1990

Six more references were obtained during the drafting of the specification.

Other Documents

Digital Dictate product brochure entitled *Turning Science Fiction into Reality*, posted on the Internet at URL <http://www.digitaldictate.com>.

Philips Natural Speech Processing product announcement entitled *A smooth-talking buy*, published in October 1996 Personal Computer World, page 36.

News article entitled *News Computing: Computing bytes*, published in September 1996 Scientific Computing World, page 10.

A white paper published by Philips Technology in September of 1994.

An article in Philips Research Topics, No. 3, September 1993, entitled *This is it*.

Apple to sell speech recognition and text-to-speech package and *New release of Kurzweil AI speech-to-text software*, both published in Speech Recognition Update, April 1996, pages 5-6.

- (b) Copies of each of the references listed above are included herein.
- (c) The following includes a detailed discussion of the references, which discussion points out with the particularity required by 37 C.F.R. § 1.111(b) and (c) how the claimed subject matter is distinguishable over the references.

The claims of the present invention recite embodiments which provide the ability to provide an interface between the output of a speech recognition engine and an application capable of processing the output. The interface enables the audio data to be played back for any output data which has been dictated, storing the character data, audio data and link data. Thus corrections may be postponed or even delegated to another person on another machine. The claims also recite embodiments which link the relationship between the output data and audio data to allow the audio data to be played back for any output data which has been dictated even if the data has a whole has been processed in such a way as to move, reorder, delete, insert or format the data.

GB2230370, *Speech Recognition Apparatus and Methods*, by Smiths Industries PLC.

This patent discloses a speech recognition method for recognising speech and outputting words with a certain level of confidence. There is no disclosure in this document of allowing the processing of the output of a speech recognition engine and the linking of the output text with audio data so that when the text is processed the links between the text and the audio data are maintained to allow playback of the audio components related to the text components.

GB2088106, *Word-processor Systems*, by The Marconi Company Ltd.

This document discloses that audio notes can be inserted into text data to allow recorded audio to be played back when the relevant portion of text is identified. This document does not disclose the text processing and link method of the present invention nor does it disclose the

method of selectively disabling the display and storage of recognised characters or the speech recognition engine for a period of time so that only the audio output of the speech recognition engine can be stored as an audio message associated with recognised text which is also stored. In GB2088106 there is no recognition of text and there is no disabling of the speech recognition or the display and storage of recognised characters.

EP0649144, *Automatic Indexing of Audio Using Speech Recognition*, by IBM Corp.

This document is concerned with the automatic indexing of audio using speech recognition. In the method disclosed the audio data is segmented into time frames which are time stamped. The corresponding text data has the time stamps associated therewith so that the text and audio can be linked. The purpose of the disclosed system is to allow an audio component recording to be indexed by selecting text in a transcript. This, as discussed on line 24 of column 1 enables faster searching for the required audio data.

This document does not disclose a system like the present invention which is concerned with allowing for the correction of mis-recognised text output from a speech recogniser by allowing a user to select and play back audio data. EP0649144 is merely concerned with indexing audio data to allow audio data to be played back.

In the present invention the text is output from the speech recognition engine and input to processing means to allow the processing of the characters e.g. word processing. This feature allows the operator to dictate text and not only perform speech recognition corrections, but also carry out amendments to the dictated text either by dictation or manually using a keyboard for example. Although the recognised characters are thus displaced or moved and other text may be inserted for which there is no corresponding audio data, the link means maintains the link between the audio data and the character component positions even after processing. Thus a user can select any of the text for play back even if that text includes text for which there is no audio data e.g. text which has been input using a keyboard.

Although at lines 31 to 39 of column 1 of EP0649144, the possibility of deleting words in the text is mentioned, since this document is merely concerned with maintaining an accurate audio recording, this document considers 'the corresponding audio segments can be moved or deleted'. It also goes on to state 'when words are inserted in the text, the corresponding audio segments can be inserted in the audio recording'. However, for application to speech recognition, such a system is impractical, if not impossible to carry out. In a speech recognition system if the digitised speech being used for recognition is output and stored, at a typical dictation speed of 80 to 100 words per minute, the stored audio data would have a file size of approximately 40 M Bytes per hour of dictation. Thus, if such data was operated upon in the manner described in EP0649144, either the processing requirements would be prohibitive or the time required for processing would be impracticable to allow for fast word processing which can be provided by the present invention. In the present invention the inventors have realised that in order to allow the maintenance of the link between audio components and recognised characters or text, if link data is formed which links the audio identifiers to the character components, the processing of the characters can be carried out at practical speeds whilst maintaining the links between the audio identifiers and the character component positions. The reason for this is that the link data can be far smaller in size e.g. typically 160 K Bytes per hour. This is some 250 times smaller than the size of the audio data file. Thus the invention is neither disclosed nor contemplated in EP0649144.

EP0077194, *Speech Recognition System*, by Sharp Kabushiki Kaisha.

This document discloses a speech recognition system which can recognise the voice of a specific speaker which has previously been analysed. This document does not disclose the output of recognised characters together with audio data in order to enable the linking of the audio identifiers to character component positions in the recognised characters.

WO93/07562, *Method and Apparatus for managing Information*, by Riverrun Technology.

This document discloses a visual display for organising and displaying speech information. This document is not at all concerned with speech recognition. Lines 33 to 36 on page 9 make it quite clear that the author considers speech recognition to be something for a 'hypothetical future'. In this system portions of the speech stream are categorised and the speech is not recognised.

US5031113, *Text-Processing System*, issued to Höllerbauer.

This document discloses a speech recognition text processing system in which the speech is recognised and the text and sound are stored. The start and end marks are allocated for words which relate in time to the sound signals stored. These marks are stored. The text can be selected and sound played to verify the accuracy of the recognition. Incorrectly recognised words can be corrected and the corrected text and original sound can be used for learning by the speech recognition engine.

This document does not disclose however the possibility of processing the text to allow the insertion, deletion and movement of text while still maintaining the link between the audio and text data. This document does not allow a user to select a passage and play back any of the audio components associated with processed text. Thus, unlike the present invention US5031113 does not allow a user to go further than merely correcting recognition errors to allow a user the flexibility to fully process a document whilst interrupting dictation.

None of the-above cited references shows the Applicants' invention as described in claims 1-45, 49-54 and 56-62 of the above-identified application. In addition, none of the above cited references, alone or in combination, teach or suggest Applicants' invention.

(d) The Applicants agree to the special examining procedures detailed in MPEP Section

PETITION TO MAKE SPECIAL
Applicant: John C. Mitchell et al.
Serial No.: 08/720,373
Docket: 491.003US1
DATA PROCESSING METHOD AND APPARATUS

708.02 VIII.

Consideration of each of the above listed references by the Examiner is respectfully requested. Pursuant to the provisions of MPEP 609, Applicants respectfully request that a copy of the attached Form 1449, marked as being considered and initialed by the Examiner, be returned to the undersigned with the next official communication.

Accordingly, the Applicants request that this Petition to Make Special be granted and the application undergo accelerated examination.


Respectfully submitted,

JOHN C. MITCHELL ET AL.

By their Attorney(s),


SCHWEGMAN, LUNDBERG, WOESSNER & KLUTH, P.A.
P.O. Box 2938
Minneapolis, MN 55402
(612) 339-0331

Date 25 March 1997
ASV/TWF/mi

By 
Ann S. Viksnins
Reg. No. 37,748

I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to Assistant Commissioner for Patents, Washington, D.C. 20231 on March 25, 1997.

Ann S. Viksnins
Name


Signature